nature immunology

SUPPLEMENTARY INFORMATION

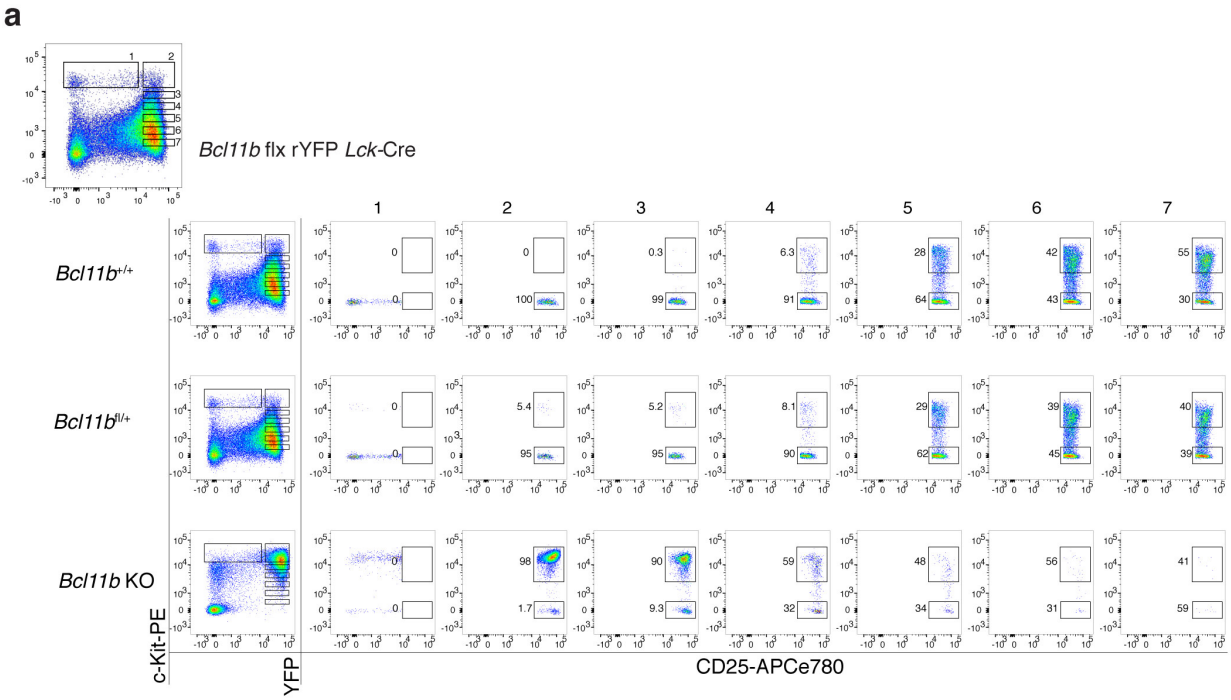
https://doi.org/10.1038/s41590-018-0238-4

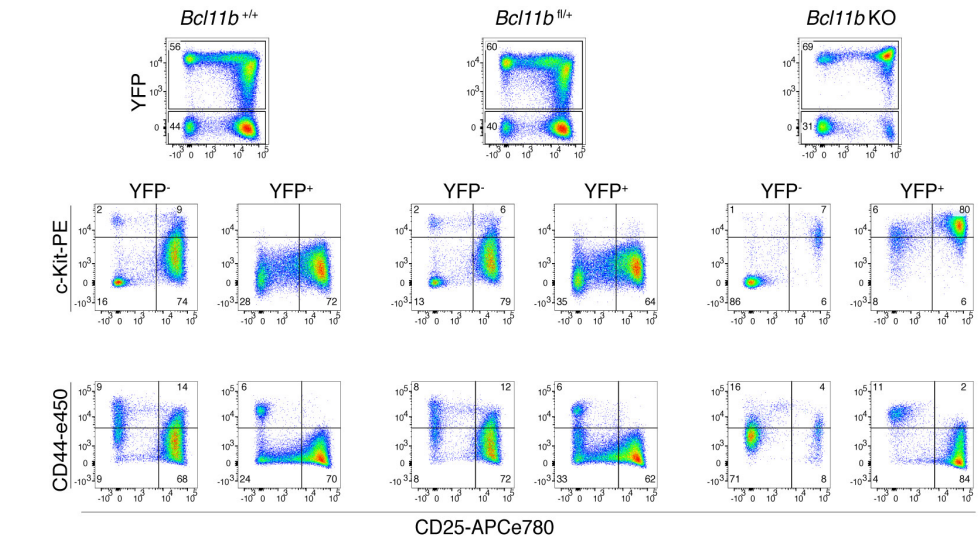
In the format provided by the authors and unedited.

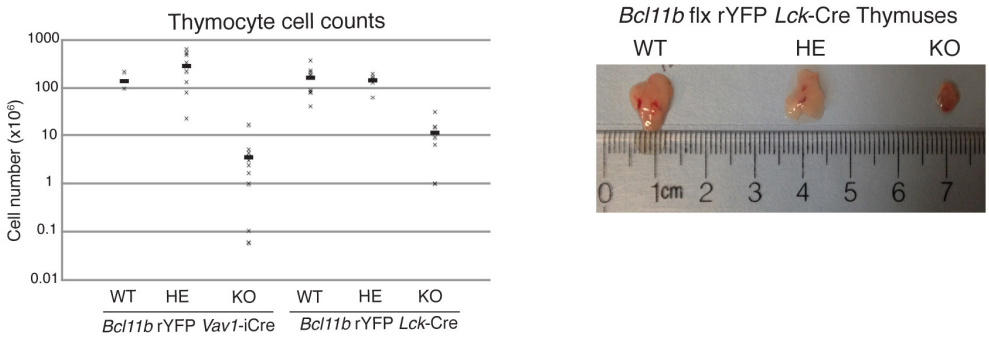
Bcl11b sets pro-T cell fate by site-specific cofactor recruitment and by repressing *Id2* and *Zbtb16*

Hiroyuki Hosokawa[1,6], Maile Romero-Wolf[1,6], Mary A. Yui[1], Jonas Ungerbäck[1,2], Maria L. G. Quiloan[1], Masaki Matsumoto[3], Keiichi I. Nakayama[3], Tomoaki Tanaka[4,5] and Ellen V. Rothenberg[1*]

[1]Division of Biology & Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [2]Division of Molecular Hematology, Lund University, Lund, Sweden. [3]Department of Molecular and Cellular Biology, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka, Japan. [4]Department of Molecular Diagnosis, Chiba University, Chuo-ku, Chiba, Japan. [5]AMED-CREST, Graduate School of Medicine, Chiba University, Chuo-ku, Chiba, Japan. [6]These authors contributed equally: Hiroyuki Hosokawa, Maile Romero-Wolf. *e-mail: evroth@its.caltech.edu


a
Bcl11b flx rYFP Lck-Cre
Bcl11b+/+
Bcl11bfl/+
Bcl11b KO
c-Kit-PE
YFP
CD25-APCe780


b
Bcl11b flx rYFP Lck-Cre
Bcl11b +/+
Bcl11b fl/+
Bcl11b KO
YFP
YFP-
YFP+
c-Kit-PE
CD44-e450
CD25-APCe780


c
Thymocyte cell counts
Cell number (x10⁶)
WT
HE
KO
Bcl11b rYFP Vav1-iCre
Bcl11b rYFP Lck-Cre
Bcl11b flx rYFP Lck-Cre Thymuses

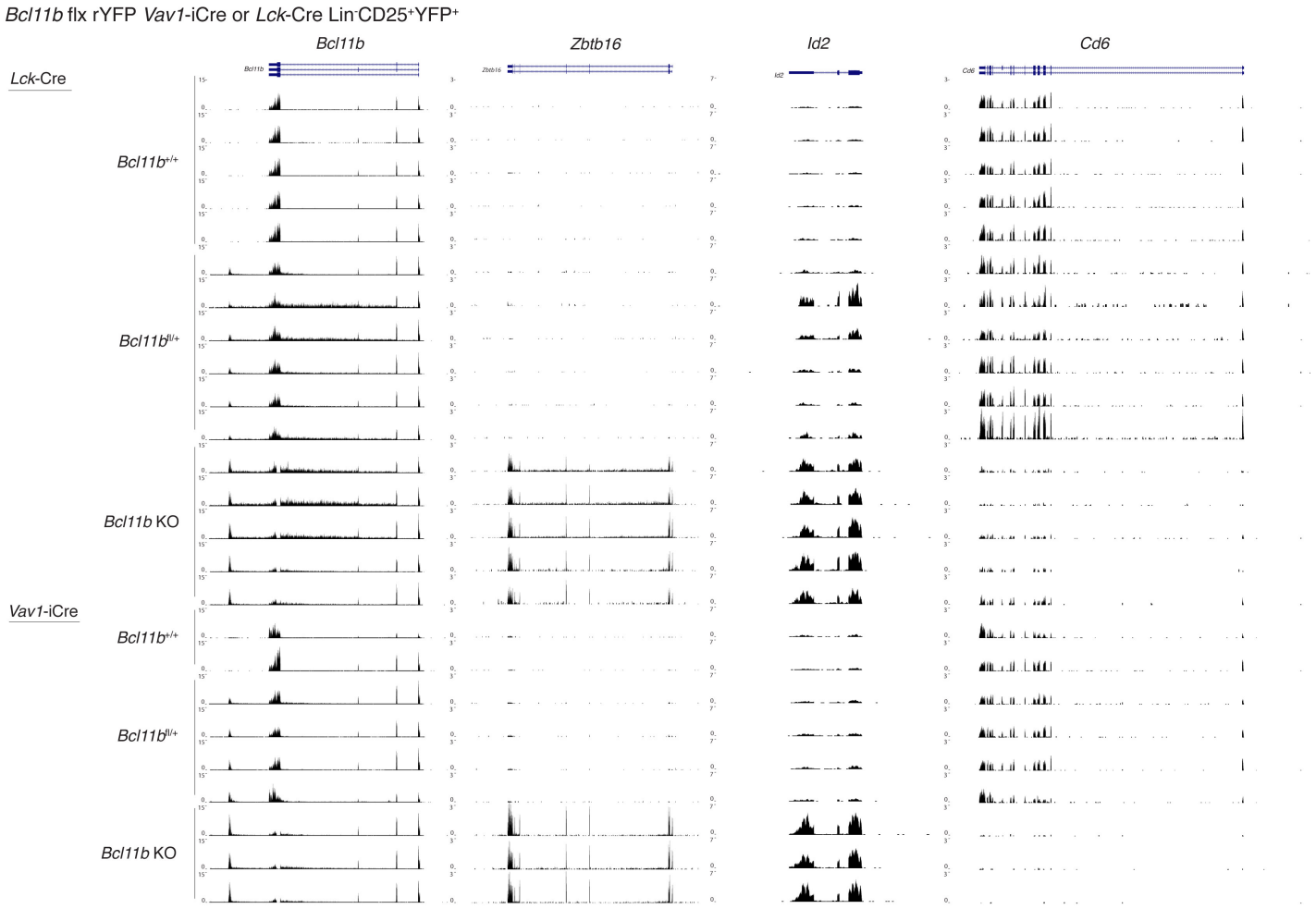
Supplementary Figure 1

Characterization of *Bcl11b* deletion systems *in vivo*

(a), Phenotypic timing of the activity of *Lck*-Cre (*Lck*-Cre^{cw}) *in vivo* based on activation of ROSA26R-YFP. For each *Bcl11b* genotype characterized in (a) and (b), results are representative of the following numbers of animals: WT, N=5; HE, N=6; KO, N=6

(b), Most CD25^{+} thymocytes in mice with *Bcl11b* deleted by *Lck*-Cre have the c-Kit^{hi+} phenotype of normal DN2a cells but the YFP^{+}CD44^{low} phenotype similar to normal DN2b/DN3a cells.

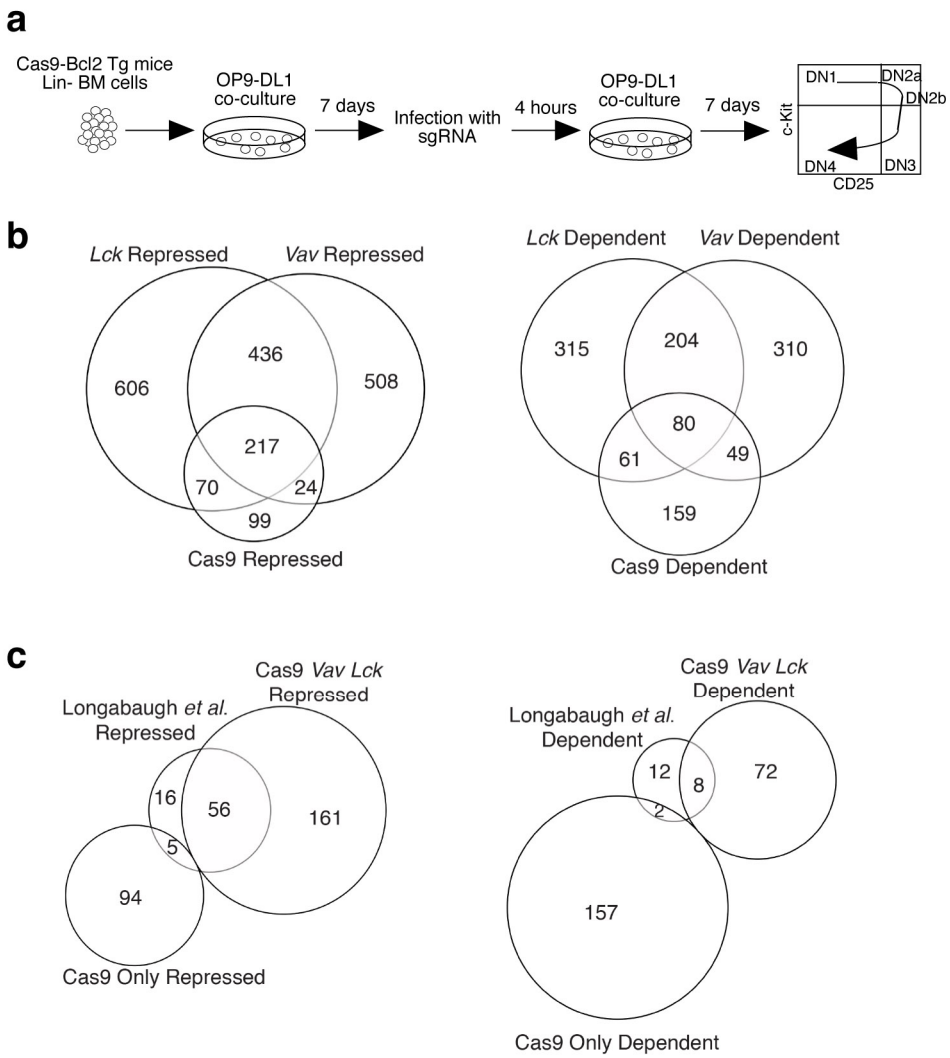
(c), (Left) *Bcl11b* deletion *in vivo* severely reduces thymocyte cellularity whether deleted before commitment by *Vav1*-iCre or after commitment by *Lck*-Cre. WT, HE, and KO refer to *Bcl11b* genotypes in the indicated Cre backgrounds: *Bcl11b*$^{+/+}$, *Bcl11b*$^{+/fl}$, and *Bcl11b*$^{fl/fl}$, respectively. Each symbol represents a different animal and the short bars indicate geometric means. N of individual animals counted for each genotype: Vav1-iCre WT N=3; Vav1-iCre HE N=10; Vav1-iCre KO N=12; Lck-Cre WT N=8; Lck-Cre HE N=5; Lck-Cre KO N=7. (Right) Representative photographs of thymuses from age-matched *Lck-Cre;ROSA26R-YFP* mice with *Bcl11b*$^{+/+}$ (WT), *Bcl11b* $^{fl/+}$ (HE), and *Bcl11b*$^{fl/fl}$ (KO) genotypes at the *Bcl11b* locus.



Supplementary Figure 2

Representative RNA-seq tracks showing *Bcl11b* deletion effects

Bcl11b, *Id2*, *Zbtb16*, and *Cd6* transcript levels are shown as determined by RNA-seq in the thymocyte samples used for the heat maps in Fig. 1b-e. All samples shown were from Lin^- (DN) $CD25^+$ cells that were positive for expression of the ROSA26R-YFP Cre reporter. Gene models (Refseq) are shown in schematics at the top. Directions of transcription are from right to left. For each gene shown, all RNA-seq tracks have y-axis adjusted to the same scale in reads/million. The deletion of the coding region in exon 4 of *Bcl11b* leads to read-though transcription and splicing to a cryptic downstream exon in these samples too, as reported previously[7].



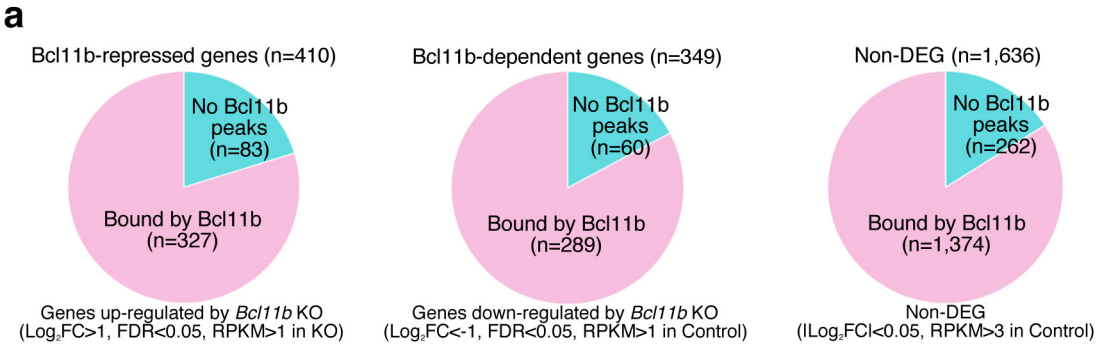
Supplementary Figure 3

Identification of consensus sets of Bcl11b target genes

(a), Experimental scheme is shown. BM-derived precursors from Cas9-Bcl2-tg mice were cultured on OP9-DL1 for 7 days, then they were infected with sgRNA. Seven days after infection, $CD25^{+}$ sgRNA transduced cells were purified and subjected to RNA-seq analysis.

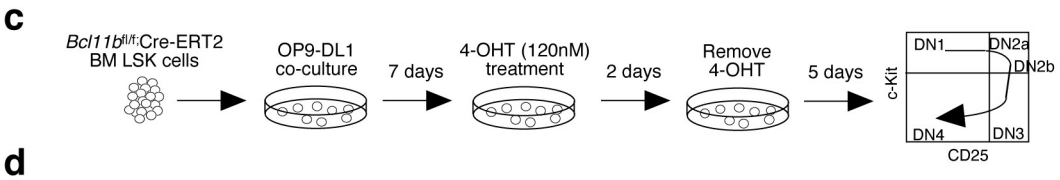
(b), Overlaps between gene sets differentially expressed when *Bcl11b* is disrupted by *Lck*-Cre or *Vav1*-iCre *in vivo,* compared with those differentially expressed when *Bcl11b* is deleted acutely by Cas9 and sgRNA in bone marrow cell precursors *in vitro*. See Supplementary Table 1 and 2 for rpkm values and gene lists. Conditions for deletion by Cas9 are described in detail below (Fig. 5) and in Methods, and rpkm values are given in Supplementary Table 3.

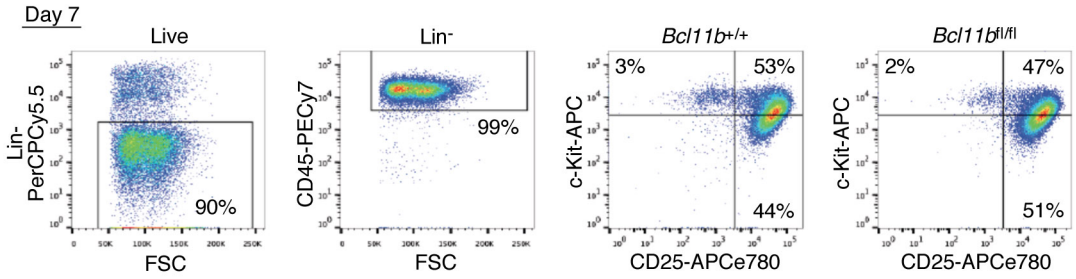
(c), Overlaps between gene sets differentially expressed by *Bcl11b* deletion in the present study with gene sets previously reported to be differentially expressed by *Bcl11b* deletion in fetal liver-derived precursors differentiating *in vitro*[7]. DEGs affected by *Lck*-Cre deletion *in vivo*, *Vav1*-iCre deletion *in vivo*, and Cas9-mediated deletion *in vitro* in the present study ("consensus set") are compared with those previously reported using retroviral Cre transduction to delete *Bcl11b* in fetal liver precursors ("Gold standard" minimal sets of responding genes in Longabaugh et al.[7]). DEG numbers in the earlier study are lower because fetal liver-derived precursors differentiate faster and the conditions used gave less consistent differentiation extents, reducing statistical significance. The consensus set includes a much higher fraction of the previously reported targets than those genes affected by Cas9-mediated deletion *in vitro* but not affected *in vivo* (a control for cell culture effects). Note that agreement among Bcl11b-repressed DEGs in all four sample types is higher than agreement among Bcl11b-dependent DEGs.

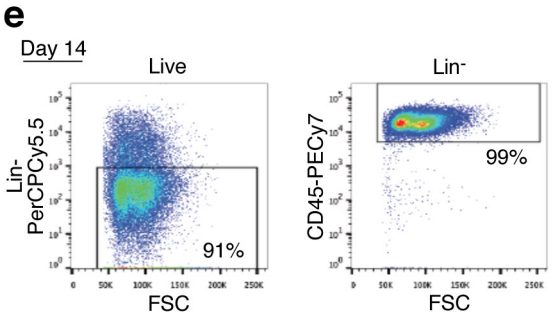


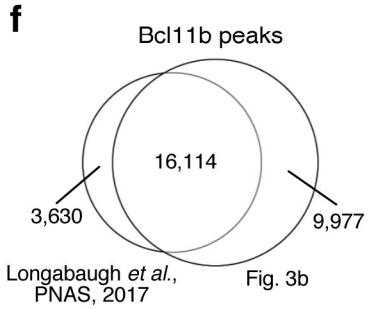
b Genes bound by Bcl11b peaks at CpG island in DEG and non-DEG

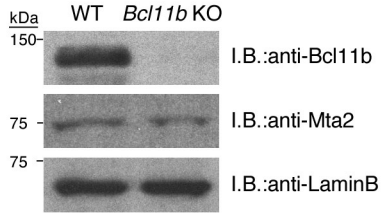
Bcl11b peaks	DEG (bound by Bcl11b) ($\|Log_2FC\|>1$, FDR<0.05, RPKM>1)	Non-DEG (bound by Bcl11b) ($\|Log_2FC\|<0.05$, RPKM>3)	*P*-value (Fisher's exact test)
CpG	98 / 616 (15.9%)	836 / 1374 (60.8%)	2.2E-16
Both	234 / 616 (38.0%)	378 / 1374 (27.5%)	9.51E-4
Non-CpG	284 / 616 (46.1%)	160 / 1374 (11.6%)	2.2E-16











Supplementary Figure 4

Functional characteristics of Bcl11b binding in pro-T cells

(a), Transcriptional responses to *Bcl11b* perturbation of genes directly bound by Bcl11b in pro-T cells. *Bcl11b*-deficient $CD25^+$ cells prepared as shown in Fig. 5a were subjected to RNA-seq analysis and differentially expressed genes (DEGs) and controls were

assessed for local Bcl11b binding by ChIP-seq. Pie charts show frequency of genes bound by Bcl11b among Bcl11b-repressed genes (left), Bcl11b -dependent genes (middle), and non-DEGs (right).

(b), Bcl11b peaks at CpG islands preferentially linked to non-DEGs in *Bcl11b*-deficient cells. The numbers of genes having Bcl11b peaks at CpG islands, non-CpG islands or both Bcl11b peaks at CpG and non-CpG islands in DEGs and non-DEGs are shown. *P* value is determined by two-sided Fisher's exact test.

(c), Tamoxifen-induced deletion of the *Bcl11b* gene in pro-T cells. Experimental scheme. BM-derived LSK cells from *Bcl11b*$^{fl/fl}$;Cre-ERT2 mice were cultured on OP9-DL1 cells for 7 days. Then, they were treated with 4-OHT (120 nM) for 2 days and cultured an additional 5 days.

(d), Flow cytometry analysis of BM-derived precursors after 7 days of OP9-DL1 culture (before 4-OHT treatment) are shown.

(e), Flow cytometric analysis of BM-derived precursors after 14 days of OP9-DL1 culture are shown.

(f), Reproducibility of Bcl11b ChIP peaks. Venn diagram shows the numbers of Bcl11b ChIP peaks in this study (Fig. 3b) and previously published data[7].

(g), Nuclear lysates from BM-derived WT and *Bcl11b*-deficient cells described in (Fig. 3a) were subjected to immunoblotting against Bcl11b, Mta2, and LaminB. Sections of western blots around the indicated size markers are shown.

Data are based on reproducible ChIP-seq peaks in two replicate samples and four replicates of RNA-seq results (a, b, f,), or are representative of two (g) or four (d, e) independent experiments.

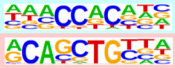
a

Bcl11b peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		Fli1 (ETS)	44.8%	14.3%	1E-3059
2		Runx1 (Runt)	36.5%	11.9%	1E-2282
3		Ptf1a (bHLH)	25.2%	13.8%	1E-521

b

Bcl11b-dependent Chd4 peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		ERG (ETS)	39.5%	7.86%	1E-611
2		Runx2 (Runt)	44.5%	14.8%	1E-403
3		Ptf1a (bHLH)	48.3%	22.2%	1E-263

Bcl11b-dependent Mta2 peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		Ets1 (ETS)	46.0%	11.9%	1E-1444
2		RUNX (Runt)	40.9%	13.4%	1E-935
3		Ptf1a (bHLH)	27.7%	15.1%	1E-209

Bcl11b-dependent Rest peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		Elk4 (ETS)	32.6%	13.1%	1E-130
2		Sp4_2	40.3%	22.8%	1E-79
3		YY1 (Zf)	9.20%	2.77%	1E-50

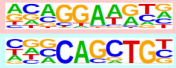
Bcl11b-dependent Ring1b peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		ERG (ETS)	45.2%	10.5%	1E-522
2		Runx1 (Runt)	24.8%	4.67%	1E-313
3		E2A (bHLH)	37.6%	15.4%	1E-196

Bcl11b-dependent LSD1 peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		ERG (ETS)	43.8%	9.29%	1E-776
2		Runx1 (Runt)	41.5%	12.8%	1E-484
3		E2A (bHLH)	46.8%	18.5%	1E-393

Bcl11b-dependent Runx1 peaks

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		Etv2 (ETS)	52.6%	14.2%	1E-975
2		E2A (bHLH)	56.3%	25.8%	1E-498
3		Runx1 (Runt)	33.8%	10.7%	1E-454

c

New Chd4 peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		RUNX (Runt)	26.1%	3.91%	1E-749
2		ERG (ETS)	25.1%	5.69%	1E-499
3		Tcf3 (HMG)	16.1%	5.19%	1E-200

New Mta2 peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		EBF	42.9%	18.0%	1E-103
2		RUNX1	31.3%	11.2%	1E-91
3		FOXA1	13.7%	2.40%	1E-82

New Rest peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		MEIS2	45.0%	2.63%	1E-151
2		Rest	14.0%	0.37%	1E-60
3		MafA (bZIP)	45.5%	18.1%	1E-32

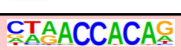
New Ring1b peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		RUNX2 (Runt)	41.7%	10.8%	1E-440
2		ERG (ETS)	16.3%	3.21%	1E-202
3		Tcf3 (HMG)	10.1%	2.53%	1E-94

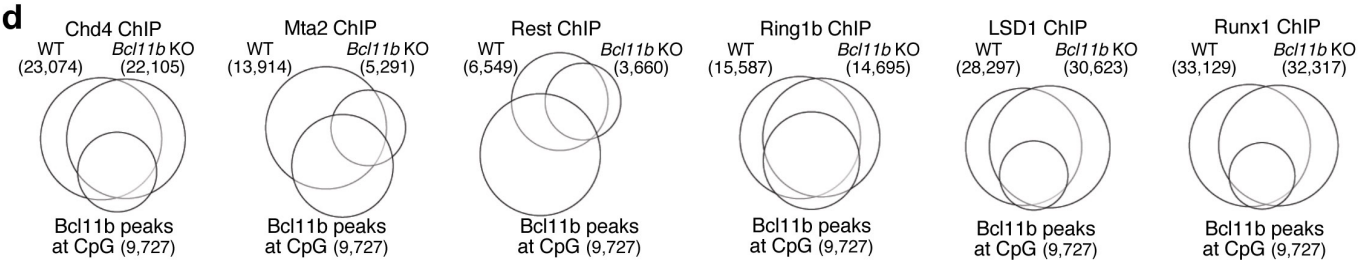
New LSD1 peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		RUNX2 (Runt)	39.9%	10.8%	1E-1179
2		ERG (ETS)	20.9%	7.92%	1E-343
3		AP-1 (bZIP)	10.3%	3.62%	1E-179

New Runx1 peaks in *Bcl11b* KO

Rank	Motif	TFs	% of Targets	% of Background	P-value
1		RUNX (Runt)	43.5%	9.38%	1E-1660
2		ERG (ETS)	18.4%	7.01%	1E-293
3		Hic1_2	30.8%	15.7%	1E-291

d



Supplementary Figure 5

Motif analysis of Bcl11b-dependent co-factor peaks

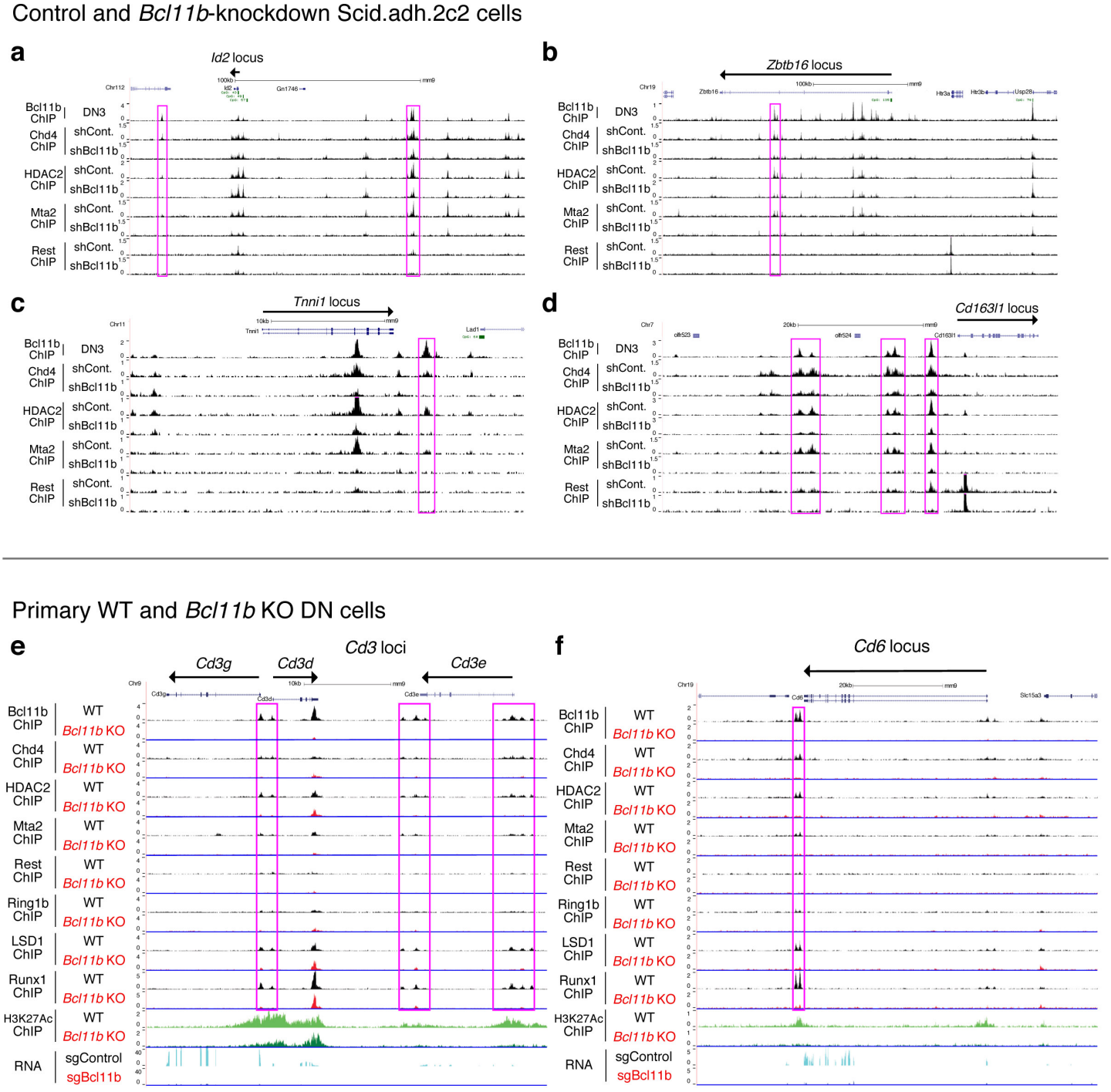
(a), Top three enriched sequence motifs of Bcl11b peaks in DN3 cells (Fig. 3b) are shown.

(b), Top three enriched sequence motifs of Bcl11b-dependent co-factor peaks in DN3 cells (Fig.3c) are shown.

(c), Top three enriched sequence motifs of new co-factor peaks in *Bcl11b*-deficient cells (Fig. 3c, 4) are shown. Note that most of these new peaks do not coincide with sites that Bcl11b normally occupies in wildtype cells: see Mta2, Rest, LSD1 and Runx1 Venn diagrams in Fig. 3c.

(d), Venn diagrams show the numbers of ChIP peaks in each category overlapping with Bcl11b peaks at CpG islands.

Data are based on reproducible ChIP-seq peaks in two independent replicate samples.



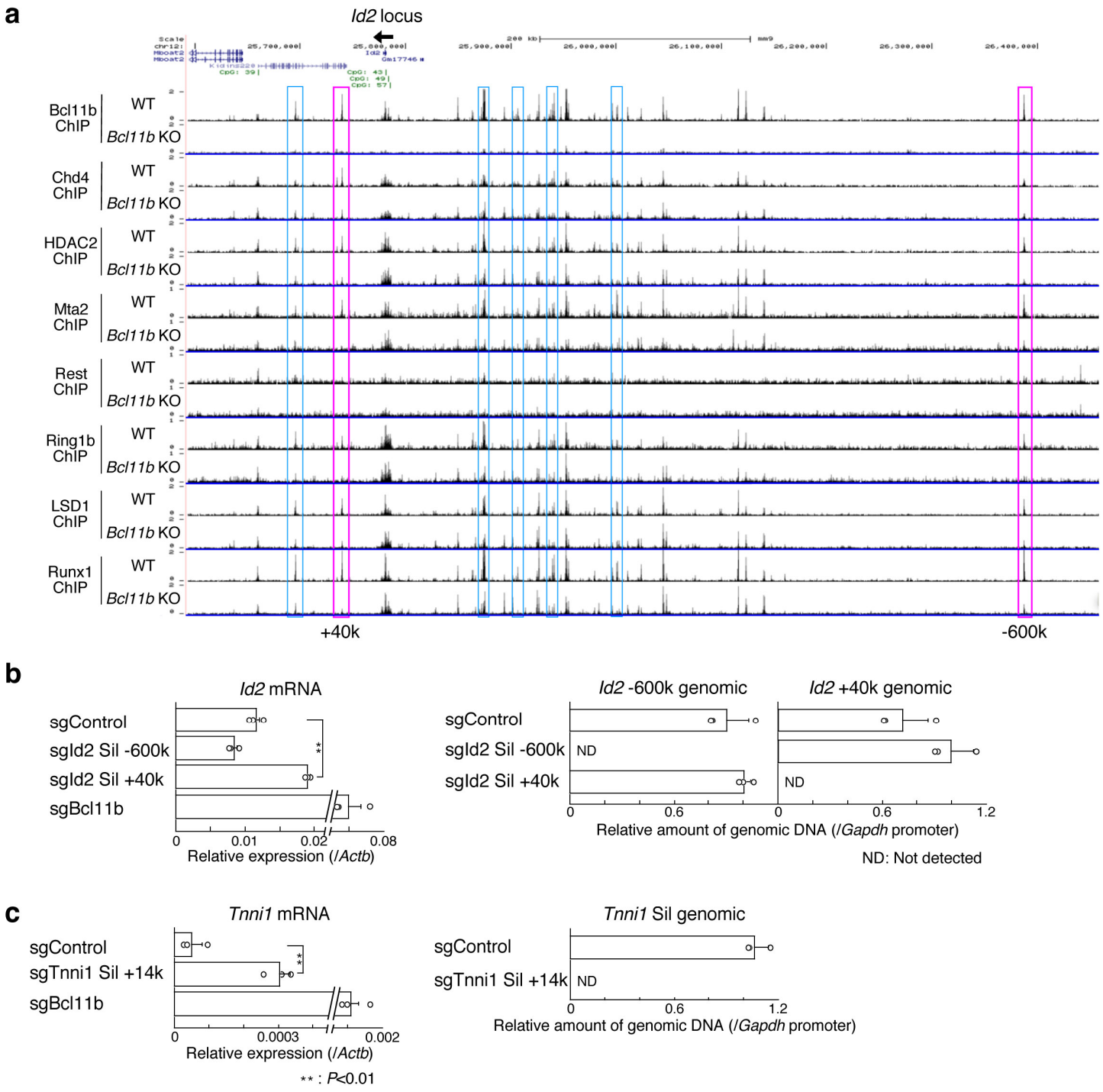
Supplementary Figure 6

Bcl11b-dependent cofactor peaks around major Bcl11b target genes

(a-d) Similar to the binding data from primary DN3 cells shown in Fig. 4a-d, this figure shows supporting data from Scid.adh.2c2 cells for binding of cofactors (Chd4, Hdac2, Mta2 and Rest) in control (shCont.) and *Bcl11b*-knockdown cells. Browser tracks show the *Id2*

(a), *Zbtb16* (b), *Tnni1* (c) and *Cd163l1* (d) loci. Sites of Bcl11b-dependent cofactor peaks in *Bcl11b*-knockdown cells are labeled with magenta rectangles. This cell-line experiment was performed once to add corroboration to the primary-cell data in Fig. 4.

(e, f), The binding patterns of Bcl11b and cofactors at the *Cd3* and *Cd6* loci in WT DN3 and *Bcl11b* KO $CD25^{+}$ cells with H3K27Ac ChIP-seq and RNA-seq tracks are shown. Bcl11b-dependent Runx1 peaks were labeled with magenta rectangles. Data are representative of two independent experiments



Supplementary Figure 7

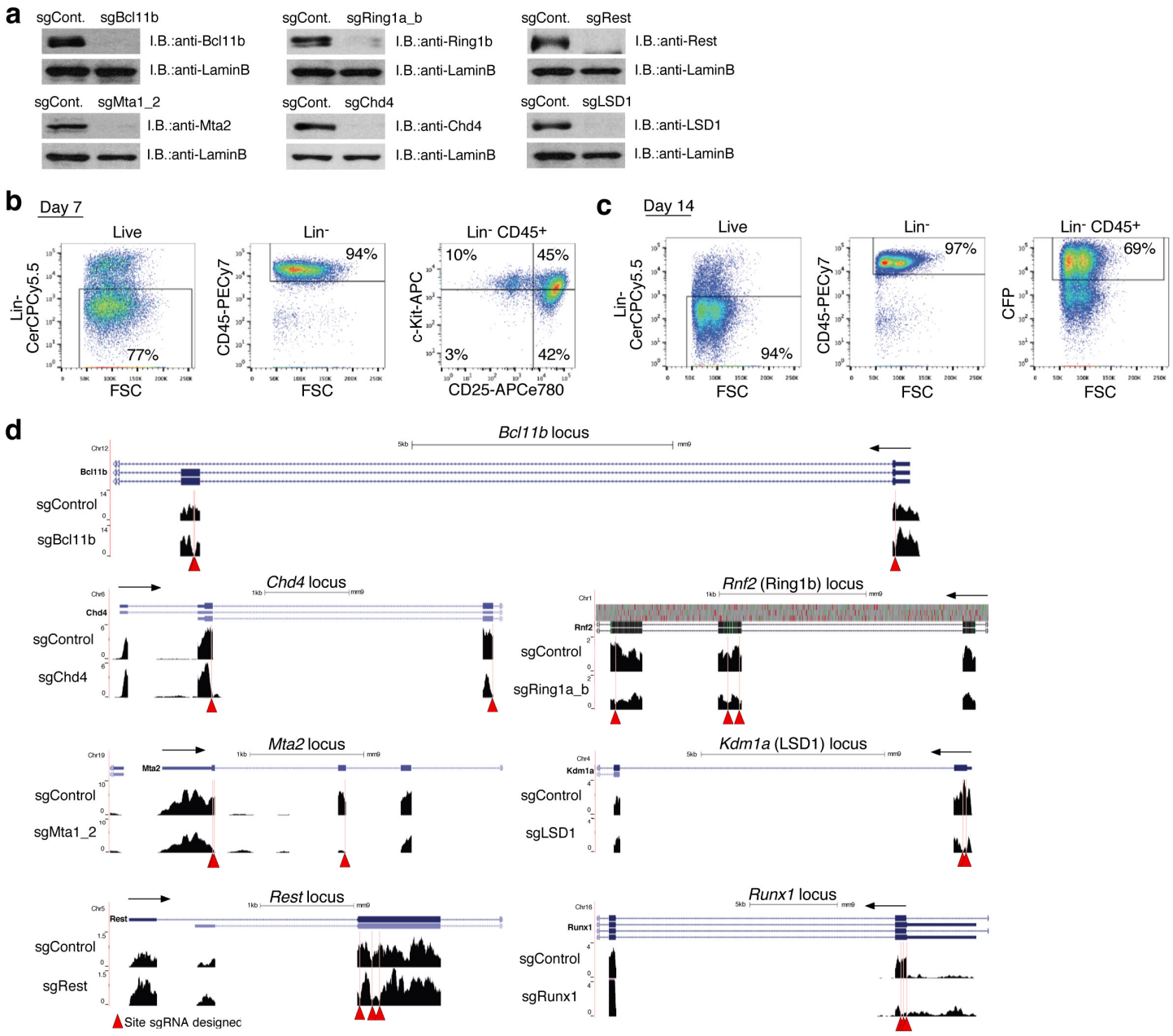
Candidate Bcl11b-dependent silencer regions for the *Id2* and *Tnni1* loci.

(a), The binding patterns of Bcl11b and cofactors at the *Id2* locus in WT DN3 and *Bcl11b*-deficient CD25^{+} cells are shown. Bcl11b-dependent cofactor peaks are labeled with magenta or light blue rectangles. We chose the two regions most clearly occupied by

Bcl11b-dependent cofactor assemblies (+40k and -600k, magenta rectangles) for further analysis.

(b, c), The genomic regions bound by Bcl11b and its cofactors can act as silencers for *Id2* and *Tnni1*. Scid.adh.2c2 cells were first infected with Cas9-GFP and GFP^+ cells were expanded. They were then transduced with sgRNA-CFP and sgRNA-hNFGR retroviruses, with the two sgRNA vector pools designed respectively against sequences on the two-sided of the targeted silencer regions. After two days of culture, CFP^+ $hNGFR^+$ infected cells were isolated by single cell sorting, and individual clones were expanded for two weeks. Genomic DNA from each clone was then isolated followed by quantitative PCR analysis for targeted genomic regions at the *Id2* (b) (magenta rectangles in (a)) and *Tnni1* (c) (magenta rectangles in Fig. 4c) loci. Each clone was also subjected to RT-qPCR analysis for transcripts from *Id2* (b) and *Tnni1* (c). The relative expression (/*Actb*) is shown. **P*<0.05, ***P*<0.01 by two-sided Student's t-test. ND, not detected.

Data are representative of two independent experiments (a) or show individual values and averages of three biological replicates with mean ±s.d. (b, c).



Supplementary Figure 8

Acute Cas9-dependent deletion of *Bcl11b* and its cofactors

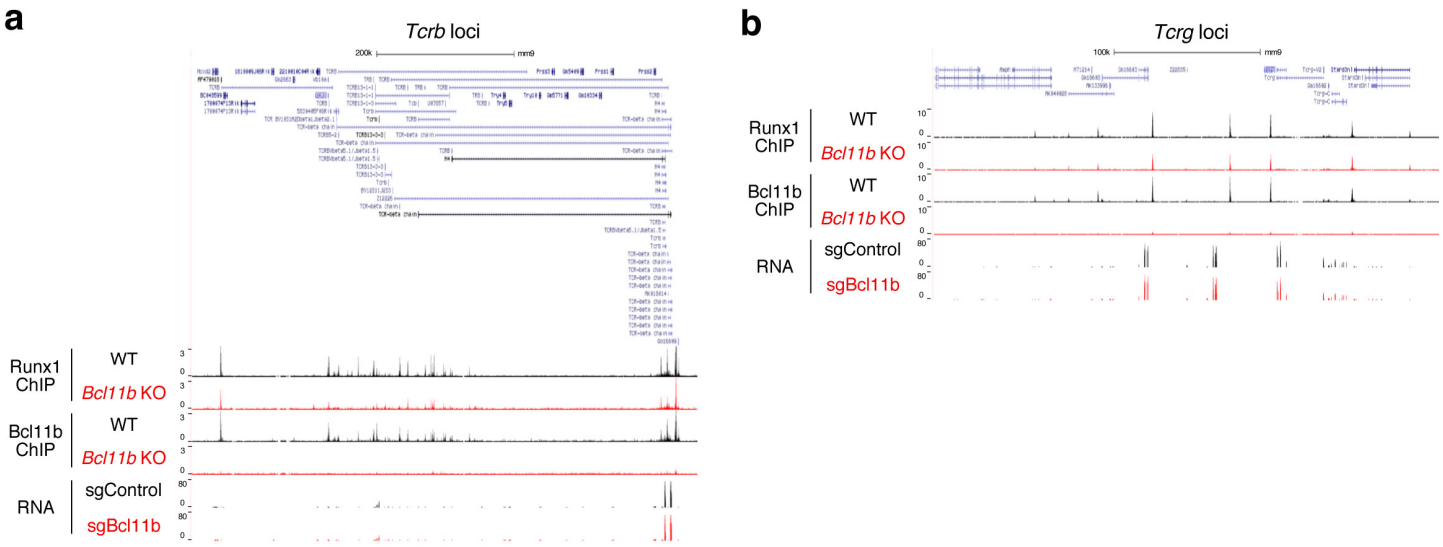
(a), Cas9-GFP introduced Scid.adh.2c2 cells were infected with sgRNA-CFP retroviruses. Two days after infection, nuclear lysates of the cells were prepared and subjected to IB with anti-Bcl11b, anti-Chd4, anti-Mta2, anti-Rest, anti-Ring1b, anti-LSD1 and anti-LaminB Abs.

(b), Flow cytometry analysis of BM-derived precursors after 7 days of OP9-DL1 culture (before retrovirus infection) is shown.

(c), Flow cytometric analysis of sgRNA transduced BM-derived pro-T cells after a total of 14 days of OP9-DL1 culture (7 d after transduction) is shown. Note the high percentage of sgRNA-transduced CFP^+ cells.

(d), Efficient disruption of Bcl11b cofactors by CRISPR-Cas9 system in primary DN cells. RNA-seq tracks for the cofactor loci are shown. Red arrowheads and dotted lines show sites against which sgRNA was designed.

Data are representative of two independent experiments (a, d) or three independent experiments (b, c).

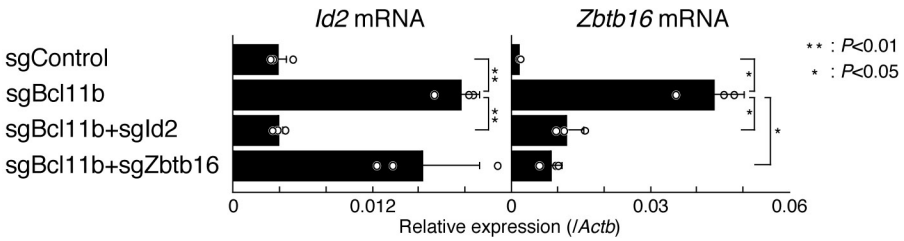


Supplementary Figure 9

Bcl11b-dependent Runx1 recruitment distinguishes the *Tcrb* and *Tcrg* genetic loci in primary pro-T cells

The binding patterns of Runx1 and Bcl11b at the *Tcrb* locus (a) and *Tcrg* loci (b) in WT DN3 and *Bcl11b* KO CD25^{+} cells are shown, with RNA-seq tracks below. Bcl11b-dependent Runx1 peaks are labeled with magenta rectangles.

Data are representative of two independent experiments.

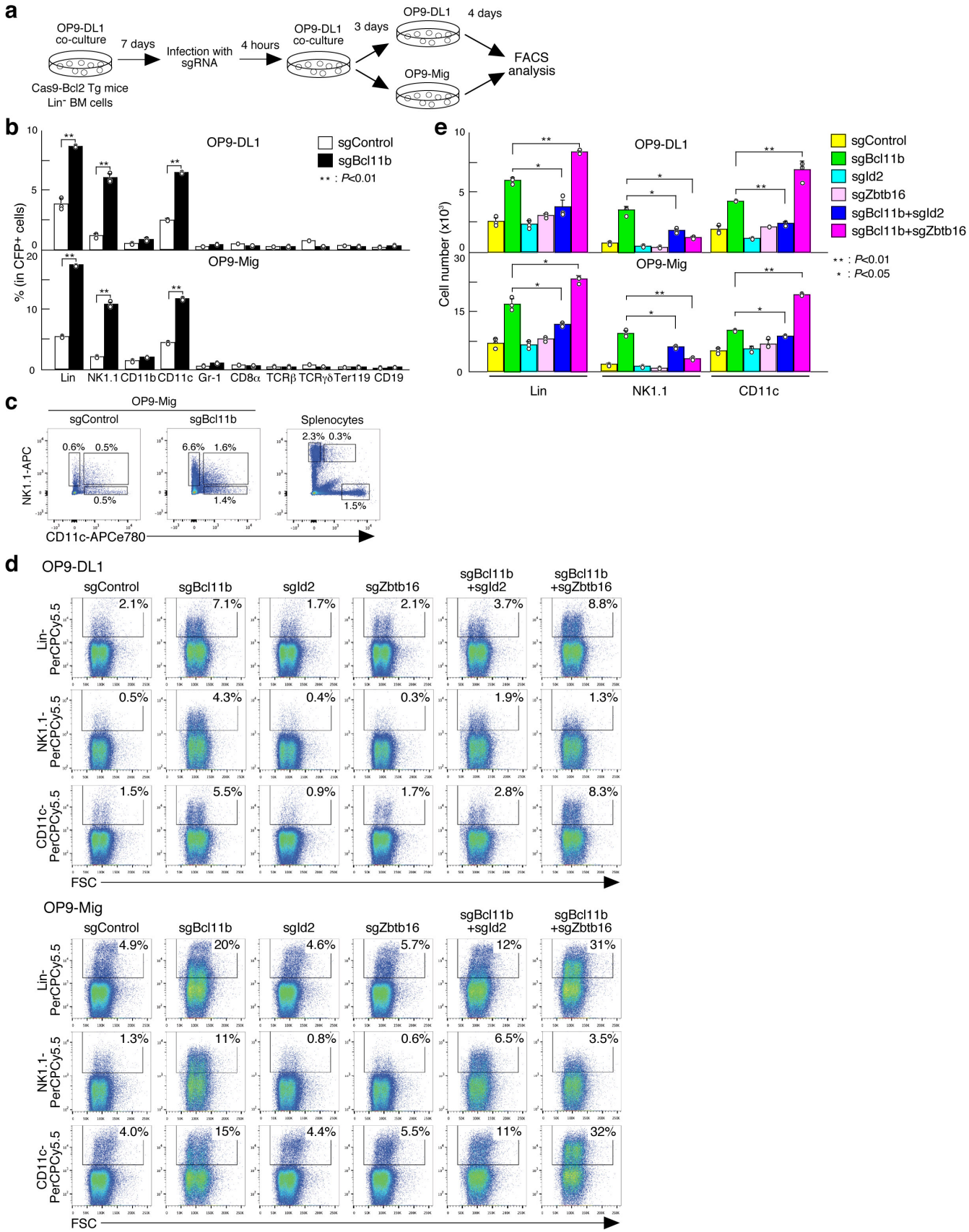


Supplementary Figure 10

Upregulation of *Zbtb16* in absence of Bcl11b depends on upregulation of *Id2*

sgRNA transduced pro-T cells cells in Fig. 6a were purified and the levels of mRNA of *Id2* and *Zbtb16* were measured by RT-qPCR. The relative expression (/*Actb*) is shown. ***P*<0.01, **P*<0.05 by two-sided Student's t-test.

Data are individual values and averages of three biological replicates with mean ±s.d..


a
OP9-DL1 co-culture
7 days
Infection with sgRNA
4 hours
OP9-DL1 co-culture
3 days
OP9-DL1
4 days
FACS analysis
OP9-Mig
Cas9-Bcl2 Tg mice Lin- BM cells
b
OP9-DL1
sgControl
sgBcl11b
** : P<0.01
OP9-Mig
% (in CFP+ cells)
Lin NK1.1 CD11b CD11c Gr-1 CD8α TCRβ TCRγδ Ter119 CD19
c
OP9-Mig
sgControl
sgBcl11b
Splenocytes
0.6% 0.5% 0.5%
6.6% 1.6% 1.4%
2.3% 0.3% 1.5%
NK1.1-APC
CD11c-APCe780
d
OP9-DL1
sgControl
sgBcl11b
sgId2
sgZbtb16
sgBcl11b +sgId2
sgBcl11b +sgZbtb16
Lin-PerCPCy5.5
NK1.1-PerCPCy5.5
CD11c-PerCPCy5.5
2.1% 7.1% 1.7% 2.1% 3.7% 8.8%
0.5% 4.3% 0.4% 0.3% 1.9% 1.3%
1.5% 5.5% 0.9% 1.7% 2.8% 8.3%
FSC
OP9-Mig
4.9% 20% 4.6% 5.7% 12% 31%
1.3% 11% 0.8% 0.6% 6.5% 3.5%
4.0% 15% 4.4% 5.5% 11% 32%
FSC
e
OP9-DL1
OP9-Mig
Cell number (x10³)
sgControl
sgBcl11b
sgId2
sgZbtb16
sgBcl11b+sgId2
sgBcl11b+sgZbtb16
** : P<0.01
* : P<0.05
Lin
NK1.1
CD11c

Supplementary Figure 11

Generation of NK1.1^{+} and CD11c^{+} cells from *Bcl11b*-deficient pro-T cells

(a), Experimental scheme is shown. BM-derived precursors were cultured on OP9-DL1 for 7 days, then they were transduced with sgRNA. Three days after transduction, at day 10 overall, they were passaged onto OP9-DL1 (upper) or OP9-Mig (lower) and cultured for 4 more days before analysis.

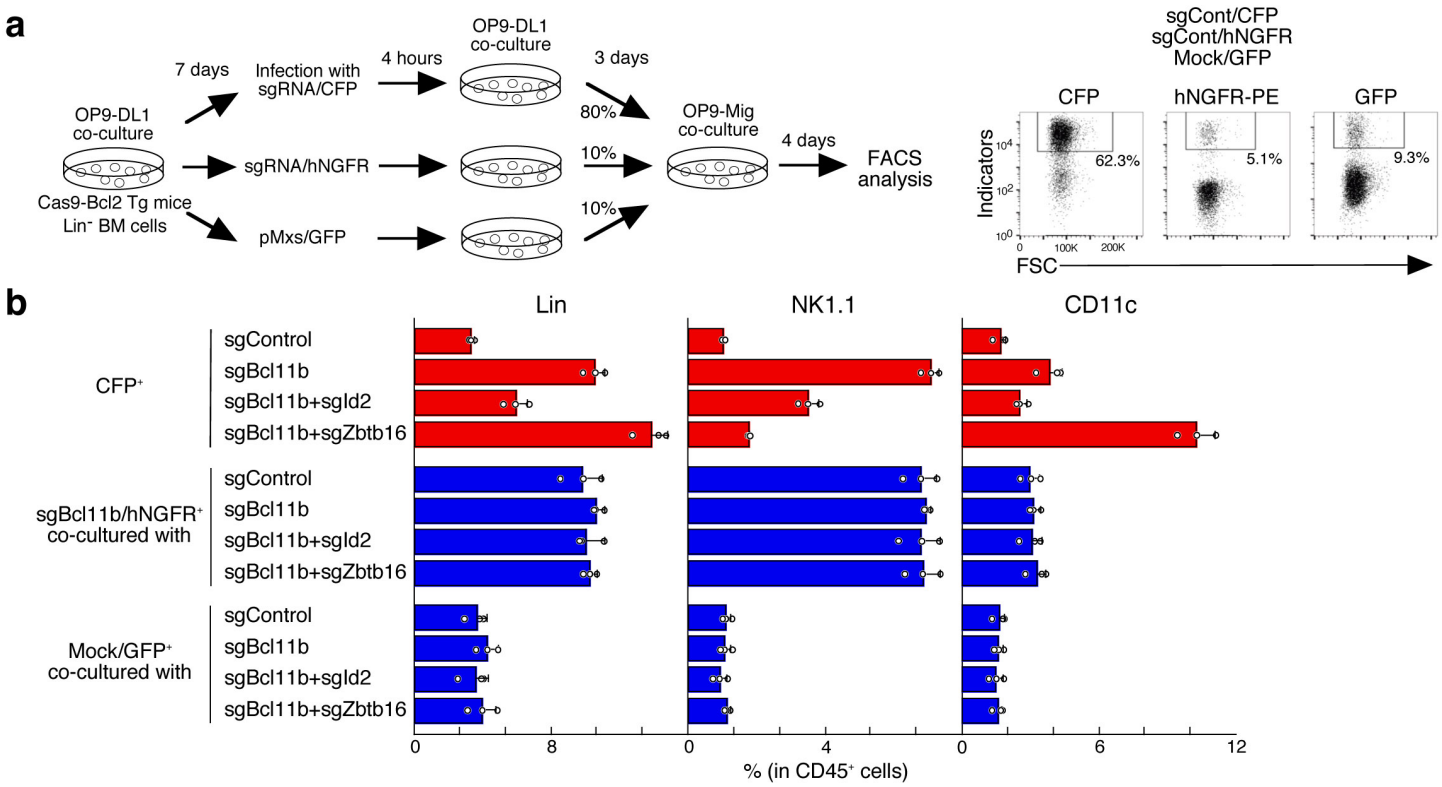
(b), Flow cytometric analyses of sgControl or sgBcl11b transduced pro- T cells were performed on day 14 (a). The percentages of Lin marker-positive cells among CFP^{+} sgRNA transduced cells are indicated. $^{**}P<0.01$ by two-sided Student's t-test. Note change of scale for OP9-Mig cocultured samples as compared to OP9-DL1 cocultured samples.

(c), Comparison of NK1.1 and CD11c expression in Bcl11b-deficient pro-T cells as compared to mature NK and DC cells in spleen. Flow cytometric analyses of sgRNA transduced BM-derived pro-T cells after 10 days of OP9-DL1 culture and 4 days of OP9-Mig culture in (a) are shown, compared with staining of fresh splenocytes from B6 mice. The percentages of NK1.1^{+}CD11c^{-}, NK1.1^{+}CD11c^{+} and NK1.1^{-}CD11c^{+} cells in CFP^{+} sgRNA transduced cells are indicated. Note differences in NK1.1 and CD11c fluorescence intensities relative to splenocytes.

(d), sgRNA transduced BM-derived precursors after 14 days of OP9-DL1 (upper) or OP9-Mig (lower) co-culture were subjected to flow cytometric analysis. Representative profiles of Forward Scatter (FSC) vs. Lin, NK1.1 or CD11c in CFP^{+} cells are shown with the percentages of cells in rectangles.

(e), Absolute cell numbers of Lin^{+}, NK1.1^{+} and CD11c^{+} in CFP^{+} sgRNA transduced cells in (d) are indicated. $^{**}P<0.01$, $^{*}P<0.05$ by two-sided Student's t-test.

Data are average of three biological replicates with mean $\pm$s.d. (b, e) or are representative of two (c) or three (d) independent experiments.



Supplementary Figure 12

Effects of Bcl11b target genes *Id2* and *Zbtb16* on lineage diversion are cell-autonomous

(a), Experimental scheme is shown. BM-derived precursors were cultured on OP9-DL1 for 7 days, then they were infected with the indicated sgRNA in a CFP vector, the indicated sgRNA in an hNGFR vector, or with pMxs-GFP empty vector, separately. Three days after infection, they were re-combined, transferred to OP9-Mig and cultured 4 more days (left). Representative FCS/CFP, hNGFR and GFP profiles on day 14 are shown with the percentages of cells in the rectangles (right).

(b) Bar graph summarizing flow cytometric analysis in Fig. 7b. Data shown are individual values and averages of three biological replicates with mean ±s.d..